Supplementary Data

Materials and methods

Mapping markers used for positional cloning of suvh4 alleles

The marker on MSH12 was amplified with the PCR primers MSH12F 5’-ACAGCTAGTCTAACC CGCGGA-3’ and MSH12R 5’-ATCCCGAC CGGAGGATACAA-3’ which generate a 419 bp product. The WS product cleaves into two approximately equal fragments with the enzyme MboII whereas the Nd-0 product is uncleaved. The marker on T21H19 was amplified with PCR primers FMDNR7F 5’-ACAGATCTAACATCTATGAG-3’ and FMDNR7R 5’-TCTTCAACCTCACCCACAC-3’, which generate an approximately 290 bp fragment from WS template DNA and an approximately 300 bp fragment from Nd-0 template DNA.

Markers used for genotyping the suvh4R302* allele and the cmt3i11a allele

To facilitate genetic analysis, a PCR-based marker was designed where the base change created by the suvh4R302* mutation is combined with a mismatch at the end of a nearby PCR primer to create a restriction site polymorphism. For this marker, the primers DPM8IF 5’-TCAGGATTTACAGTGATAAGTACCGACTGAAT-3’ and DPM8IR 5’-TGGAACTTTGGTATATGGAACG-3’, where the underlined base indicates a mismatch with the genomic sequence, were used to amplify a 144 bp fragment from the SUVH4 gene. The fragment amplified from wild type SUVH4 template DNA cleaves with ClaI into 111 bp and 33 bp products, whereas the fragment amplified from suvh4R302* template DNA is uncleaved by ClaI.

The cmt3illa allele, which creates an MseI site, was scored by amplifying the relevant region of the CMT3 gene with the primers cmt3i11F 5’-GTTCTGCGTCTAGTTAATTG-3’ and cmt3i11R 5’-GTGACCACGTGTTCCCTTGCG-3’. These primers amplify a 150 bp fragment. In wild type CMT3, the fragment cleaves once with MseI into 140 bp and 10 bp fragments. In the cmt3illa mutant, the fragment cleaves twice with MseI into fragments of 85, 55, and 10 bp.
Figure 1. *HincII* Southern blot assays for non-CG PAI1-PAI4 and PAI2 methylation in *lhp1*, *suvh4*, and *cmt3* mutant strains. (A) The *HincII* (Hc) restriction map of each WS PAI locus is shown, with the probed regions indicated by gray bars and the bisulfite sequenced regions indicated by hatched bars. The sequences immediately flanking each *HincII* site tested are shown, indicating the non-CG contexts of the outer cytosines. These outer cytosines are the positions at which the enzyme is sensitive to methylation. Note that the divergent *PAI3* gene does not carry an internal *HincII* site and is thus not informative for methylation status by this assay. (B) Genomic DNAs prepared from four-week-old plants of the indicated genotypes were cleaved with *HincII* and used for Southern blot analysis with a *PAI* probe. P1-P4 is *PAI1-PAI4*, P2 is *PAI2*, and P3 is *PAI3*, with asterisks indicating the positions of species methylated at internal *HincII* sites.
**Figure 2.** Bisulfite sequencing analysis of the PAI genes in wild type WS, *suvh4, cmt3,* and *suvh4 cmt3* mutant backgrounds. Bisulfite genomic methylation sequencing was performed for the top strands of the *PAI1* and *PAI2* upstream regions in wild type WS, WS *pai1 cmt3G456D*, WS *pai1 suvh4R302* or WS *pai1 suvh4R302* *cmt3illa* DNA. For each region, eight independent molecules were sequenced. Vertical lines indicate positions of cytosines, with the height of each line representing how many sequenced molecules had 5-methyl-cytosine (5-MeC). Black indicates CG cytosines, blue indicates CNG cytosines, and red indicates other cytosines. Asterisks indicate sites with no methylation. The black horizontal line indicates the region of PAI identity, and the gray horizontal line indicates flanking upstream heterologous sequence unique to each gene. The exon and intron structures of *PAI1* and *PAI2* are shown as open boxes and dashed lines respectively under each sequence. These structures are based on full-length cDNA sequences for each gene (Melquist *et al.* 1999). Data for wild type WS and WS *pai1 cmt3* are reproduced from previous publications (Luff *et al.*, 1999; Bartee *et al.*, 2001).
Figure 3. DNA sequence alterations in seven suvh4 alleles. The coding region of the WS SUVH4 gene is shown, with exons in uppercase letters, and introns in lower case letters. The positions of bases affected by mutation are indicated by underlining. The mutant base alteration for each allele is shown above the sequence.
Figure 4. Two splice junction *suvh4* mutations lead to a variety of splicing defects. Exons are denoted by open boxes and introns are denoted by lines for the 5’ end of the *SUVH4* coding region. The positions of the intron 1 acceptor site mutation *suvh4i1a* and the intron 3 acceptor site mutation *suvh4i3a* are marked with asterisks. For each mutant, the splice variants recovered by cloning and sequencing independent reverse transcriptase-PCR products are shown, with a brief description of the consequences of the mis-splicing for the protein-coding region.